GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 21, 2001, 10:01:49 ; Search time 2052.79 Seconds
                                          (without alignments)
                                          12935.110 Million cell updates/sec

Title: US-09-587-111-4
Perfect score: 2809
Sequence: 1 ggctagcctgtcctgacagg..........aaaaaaaaaaaaaaaaaaa 2809

Scoring table: IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0

Searched: 10228115 seqs, 4726426750 residues

Total number of hits satisfying chosen parameters: 20456230

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :

EST:*

1: gb_est1:*	44: em_esthum10:*			
2: gb_est2:*	45: em_esthum11:*			
3: gb_est3:*	46: em_esthum12:*			
4: gb_est4:*	47: em_esthum13:*			
5: gb_est5:*	48: em_esthum14:*			
6: gb_est6:*	49: em_esthum15:*			
7: gb_est7:*	50: em_esthum16:*			
8: gb_est8:*	51: em_esthum17:*			
9: gb_est9:*	52: em_esthum18:*			
10: gb_est10:*	53: em_esthum19:*			
11: gb_est11:*	54: em_esthum20:*			
12: gb_est12:*	55: em_esthum21:*			
13: gb_est13:*	56: em_esthum22:*			
14: gb_est14:*	57: em_esthum23:*			
15: gb_est15:*	58: em_esthum24:*			
16: gb_est16:*	59: em_esthum25:*			
17: gb_est17:*	60: em_esthum26:*			
18: gb_est18:*	61: em_esthum27:*			
19: gb_est19:*	62: em_esthum28:*			
20: gb_est20:*	63: em_estin1:*			
21: gb_est21:*	64: em_estin2:*			
22: gb_est22:*	65: em_estin3:*			
23: gb_est23:*	66: em_estin4:*			
24: gb_est24:*	67: em_estin5:*			
25: gb_est33:*	68: em_estom1:*			
26: gb_est34:*	69: em_estom2:*			
27: gb_est35:*	70: em_estov1:*			
28: gb_est36:*	71: em_estov2:*			
29: gb_est37:*	72: em_estpl1:*			
30: gb_est38:*	73: em_estpl2:*			
31: gb_est39:*	74: em_estpl3:*			
32: gb_est40:*	75: em_estpl4:*			
33: em_estba:*	76: em_estpl5:*			
34: em_estfun:*	77: em_estpl6:*			
35: em_esthum1:*	78: em_estpl7:*			
36: em_esthum2:*	79: em_estpl8:*			
37: em_esthum3:*	80: em_estpl9:*			
38: em_esthum4:*	81: em_estpl10:*			
39: em_esthum5:*	82: em_estro1:*			
40: em_esthum6:*	83: em_estro2:*			
41: em_esthum7:*	84: em_estro3:*			
42: em_esthum8:*	85: em_estro4:*			
43: em_esthum9:*	86: em_estro5:*			

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 21, 2001, 10:01:49 ; Search time 2052.79 Seconds
                                          (without alignments)
                                          12935.110 Million cell updates/sec

Title: US-09-587-111-4
Perfect score: 2809
Sequence: 1 ggctagcctgtcctgacagg..........aaaaaaaaaaaaaaaaaaa 2809

Scoring table: IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0

Searched: 10228115 seqs, 4726426750 residues

Total number of hits satisfying chosen parameters: 20456230

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :

EST:*

Col 1	Col 2	Col 3
1: gb_est1:*	44: em_esthum10:*	87: em_estro6:*
2: gb_est2:*	45: em_esthum11:*	88: em_estro7:*
3: gb_est3:*	46: em_esthum12:*	89: em_estro8:*
4: gb_est4:*	47: em_esthum13:*	90: em_estro9:*
5: gb_est5:*	48: em_esthum14:*	91: em_estro10:*
6: gb_est6:*	49: em_esthum15:*	92: em_estro11:*
7: gb_est7:*	50: em_esthum16:*	93: em_estro12:*
8: gb_est8:*	51: em_esthum17:*	94: em_estro13:*
9: gb_est9:*	52: em_esthum18:*	95: em_estro14:*
10: gb_est10:*	53: em_esthum19:*	96: em_estro15:*
11: gb_est11:*	54: em_esthum20:*	97: em_estro16:*
12: gb_est12:*	55: em_esthum21:*	98: em_estro17:*
13: gb_est13:*	56: em_esthum22:*	99: em_estro18:*
14: gb_est14:*	57: em_esthum23:*	100: em_estro19:*
15: gb_est15:*	58: em_esthum24:*	101: em_estro20:*
16: gb_est16:*	59: em_esthum25:*	102: gb_est25:*
17: gb_est17:*	60: em_esthum26:*	103: gb_est26:*
18: gb_est18:*	61: em_esthum27:*	104: gb_est27:*
19: gb_est19:*	62: em_esthum28:*	105: gb_est28:*
20: gb_est20:*	63: em_estin1:*	106: gb_est29:*
21: gb_est21:*	64: em_estin2:*	107: gb_est30:*
22: gb_est22:*	65: em_estin3:*	108: gb_est31:*
23: gb_est23:*	66: em_estin4:*	109: gb_est32:*
24: gb_est24:*	67: em_estin5:*	110: gb_est41:*
25: gb_est33:*	68: em_estom1:*	111: gb_est42:*
26: gb_est34:*	69: em_estom2:*	112: gb_est43:*
27: gb_est35:*	70: em_estov1:*	113: gb_est44:*
28: gb_est36:*	71: em_estov2:*	114: gb_est45:*
29: gb_est37:*	72: em_estpl1:*	115: gb_est46:*
30: gb_est38:*	73: em_estpl2:*	116: gb_est47:*
31: gb_est39:*	74: em_estpl3:*	
32: gb_est40:*	75: em_estpl4:*	
33: em_estba:*	76: em_estpl5:*	
34: em_estfun:*	77: em_estpl6:*	
35: em_esthum1:*	78: em_estpl7:*	
36: em_esthum2:*	79: em_estpl8:*	
37: em_esthum3:*	80: em_estpl9:*	
38: em_esthum4:*	81: em_estpl10:*	
39: em_esthum5:*	82: em_estro1:*	
40: em_esthum6:*	83: em_estro2:*	
41: em_esthum7:*	84: em_estro3:*	
42: em_esthum8:*	85: em_estro4:*	
43: em_esthum9:*	86: em_estro5:*	

117: gb_est48:*
118: gb_est49:*
119: gb_est50:*
120: gb_est51:*
121: gb_est52:*
122: gb_est53:*
123: gb_est54:*
124: gb_est55:*
125: gb_est56:*
126: gb_est57:*
127: gb_est58:*
128: gb_est59:*
129: gb_est60:*
130: gb_est61:*
131: gb_est62:*
132: gb_est63:*
133: gb_est64:*
134: gb_est65:*
135: gb_est66:*
136: gb_est75:*
137: gb_est76:*
138: gb_est77:*
139: gb_est78:*
140: gb_est79:*
141: gb_est80:*
142: gb_est81:*
143: gb_est82:*
144: gb_est83:*
145: gb_est84:*
146: gb_est85:*
147: gb_est86:*
148: gb_est87:*
149: gb_est88:*
150: gb_est89:*
151: gb_est90:*
152: gb_est99:*
153: gb_est100:*
154: gb_est101:*
155: gb_est102:*
156: gb_est103:*
157: gb_est104:*
158: gb_est105:*
159: gb_est106:*
160: gb_est67:*
161: gb_est68:*
162: gb_est69:*
163: gb_est70:*
164: gb_est71:*
165: gb_est72:*
166: gb_est73:*
167: gb_est74:*
168: gb_est91:*
169: gb_est92:*
170: gb_est93:*
171: gb_est94:*
172: gb_est95:*
173: gb_est96:*
174: gb_est97:*
175: gb_est98:*
176: em_esthum29:*
177: em_esthum30:*
178: em_esthum31:*
179: em_esthum32:*
180: em_esthum33:*
181: em_estom3:*
182: em_estpl11:*
183: em_estro21:*
184: em_estro22:*
185: em_estro23:*
186: em_htc:*
187: gb_est107:*
188: gb_est108:*
189: gb_est109:*

190: gb_est110:*
191: gb_est111:*
192: gb_htc:*
193: em_gss_fun:*
194: em_gss_hum1:*
195: em_gss_hum2:*
196: em_gss_hum3:*
197: em_gss_hum4:*
198: em_gss_hum5:*
199: em_gss_hum6:*
200: em_gss_hum7:*
201: em_gss_hum8:*
202: em_gss_hum9:*
203: em_gss_inv1:*
204: em_gss_inv2:*
205: em_gss_inv3:*
206: em_gss_Other:*
207: em_gss_pln1:*
208: em_gss_pln2:*
209: em_gss_pro:*
210: em_gss_rod1:*
211: em_gss_rod2:*
212: em_gss_rod3:*
213: em_gss_rod4:*
214: em_gss_rod5:*
215: em_gss_vrt1:*
216: em_gss_vrt2:*
217: em_gss_vrt3:*
218: gb_gss1:*
219: gb_gss2:*
220: gb_gss3:*
221: gb_gss4:*
222: gb_gss5:*
223: gb_gss6:*
224: gb_gss7:*
225: gb_gss8:*
226: gb_gss9:*
227: gb_gss10:*
228: gb_gss11:*
229: gb_gss12:*
230: gb_gss13:*
231: gb_gss14:*
232: gb_gss15:*
233: gb_gss16:*
234: gb_gss17:*
235: gb_gss18:*
236: gb_gss19:*
237: gb_gss20:*
238: gb_gss21:*
239: gb_gss22:*
240: gb_gss23:*
241: gb_gss24:*
242: gb_gss25:*
243: gb_gss26:*
244: gb_gss27:*
245: gb_gss28:*
246: gb_gss29:*
247: gb_gss30:*
248: gb_gss31:*
249: gb_gss32:*
250: gb_gss33:*
251: gb_gss34:*
252: em_gss_inv4:*
253: em_gss_rod6:*
254: em_gss_rod7:*
255: em_gss_rod8:*
256: gb_gss35:*
257: gb_gss36:*
258: gb_gss37:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

RESULT 1

LOCUS   AK008689   2270 bp   mRNA   HTC   08-FEB-2001
DEFINITION Mus musculus adult male stomach cDNA, RIKEN full-length enriched
            library, clone:2210009M20, full insert sequence.
ACCESSION   AK008689
VERSION   AK008689.1 GI:12843037
KEYWORDS   HTC
SOURCE   Mus musculus (strain:C57BL/6J) adult male stomach cDNA to mRNA,
            clone_lib:RIKEN full-length enriched mouse cDNA library
            clone:2210009M20.
            CAP trapper.
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (sites)
AUTHORS   Carninci,P. and Hayashizaki,Y.
TITLE   High-efficiency full-length cDNA cloning

REFERENCE   2 (sites)
AUTHORS   Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
            Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE   Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL   Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE   20499374
REFERENCE   3 (sites)
AUTHORS   Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
            Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,
            Kikuchi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T.,
            Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T.,
            Kashiwagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
            Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T.,
            Matsuura,S., Okazaki,Y., Muramatsu,M., Inoue,Y. and Hayashizaki,Y.
TITLE   RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer
JOURNAL   Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE   20530913
REFERENCE   4 (sites)
AUTHORS   Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A.,
            Arakawa,T., Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M.,
            Hanagaki,T., Hara,A., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,F.,
            Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kato,H., Kawai,J.,
            Kojima,Y., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T.,
            Miyazaki,A., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Okazaki,Y.,
            Okido,T., Owa,C., Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H.,
            Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
            Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F.,
            Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yasunishi,A.,
            Yoshida,K., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
TITLE   Functional annotation of a full-length mouse cDNA collection
JOURNAL   Nature 409, 685-690 (2001)
REFERENCE   5 (bases 1 to 2270)
AUTHORS   FANTOM Consortium.
TITLE   The RIKEN Genome Exploration Research Group Phase II Team and
            FANTOM Consortium.
JOURNAL   Direct Submission
            Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of
            Physical and Chemical Research (RIKEN), Laboratory for Genome
            Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
            RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
            Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp;
            URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
            Fax:81-45-503-9216)
COMMENT   Please visit our web site (http://genome.gsc.riken.go.jp/) for
            further details.
            cDNA library was prepared and sequenced in Mouse Genome
            Encyclopedia Project of Genome Exploration Research Group in Riken
            Genomic Sciences Center and Genome Science Laboratory in RIKEN.
            Division of Experimental Animal Research in Riken contributed to
            prepare mouse tissues. First strand cDNA was primed with a primer
            [5' GAGAGAGAGAAGGATCCAAGAGCTCTGTTTTTTTTTTTTTTTTTTTTVN 3'], cDNA was
            prepared by using trehalose thermo-activated reverse transcriptase
            and subsequently enriched for full-length by cap-trapper. cDNA went
            through one round of normalization to Rot = 5.0 and subtraction by
            Rot = 20.0. Second strand cDNA was prepared with the primer adapter
            of sequence [5'
            GAGAGAGAGATTCTCGAGTTAATTAAATTAAATCCCCCCCCCCCCC 3']. cDNA was cleaved
            with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI.
            Host: SOLR.
FEATURES   Location/Qualifiers
     source   1..2270
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /db_xref="MGD:MGI:1341836"
                     /db_xref="MGD:MGI:1905060"
                     /clone="2210009M20"
                     /sex="male"
                     /tissue_type="stomach"
                     /clone_lib="RIKEN full-length enriched mouse cDNA library"
                     /dev_stage="adult"

SUMMARIES

Result No.	Score	% Query Match	Length	DB ID	Description
1	1505.8	53.6	2270	192	AK008689	AK008689 Mus muscu
2	755.8	26.9	793	108	AU141855	AU141855
3	651.4	23.2	953	175	BG254423	BG254423 602359021
4	647.8	23.1	840	108	BF689928	BF689928 602186482
5	629.8	22.4	686	123	AW976725	AW976725 EST388834
6	618.4	22.0	841	108	AU136431	AU136431 AU136431
7	610.2	21.7	682	102	AI801897	AI801897 tx28h02.x
8	596.2	21.2	682	168	BF732920	BF732920 nae17b01.
9	588	20.9	588	171	BF940288	BF940288 7o43f06.x
10	554	19.7	984	168	BF690154	BF690154 602186482
11	536.0	19.1	543	148	BF433303	BF433303 7q60f10.x
12	532.4	19.0	550	145	BF195711	BF195711 7n87e08.x
13	522.8	18.6	536	117	AW517339	AW517339 xp93e02.x
14	522.2	18.6	536	138	BE670695	BE670695 7e38h05.x
15	513	18.3	148	148	BF436096	BF436096 nab77a02-
16	510.8	18.2	514	113	AW205776	AW205776 UI-H-BI1-
17	509.8	18.1	516	165	BE247412	BE247412 TCBAP1E63
18	508.8	18.1	512	116	AW467012	AW467012 ha07e10.x
19	505.6	18.0	513	143	BF058747	BF058747 7k34c09.x
20	503.4	17.9	987	168	BF689977	BF689977 602186351
21	480	17.1	480	16	AI126236	AI126236 qc55a11.x
22	475.2	16.9	481	14	AA994465	AA994465 ou38h02.s
23	474.2	16.9	496	159	N29128	N29128 yx47e06.r1
24	463.6	16.5	470	165	BE438897	BE438897 TCBAP1E15
25	459.8	16.4	463	114	AW303365	AW303365 xv16c01.x
26	458.4	16.3	460	18	AI305201	AI305201 q190c09.x
27	454.2	16.2	486	151	BF594605	BF594605 7o53e04.x
28	454	16.2	486	11	AA741232	AA741232 ny96a11.s
29	453.4	16.2	457	19	AI373010	AI373010 q202f09.x
30	452.6	16.1	459	22	AI633879	AI633879 tt04f03.x
31	451.8	16.1	462	168	BE466577	BE466577 hz22f04.x
32	450.8	16.0	811	168	BF724464	BF724464 bx05b12.y
33	450.2	16.0	168	168	BF690051	BF690051 602186351
34	447.2	15.9	544	103	AI870136	AI870136 wk98c05.x
35	447	15.9	461	190	W44731	W44731 zb98d02.s1
36	445.2	15.8	475	104	AI951557	AI951557 wv36f05.x
37	444	15.8	455	18	AI268256	AI268256 q03c06.x
38	435.0	15.5	439	24	AI738529	AI738529 wi33b04.x
39	431.8	15.4	435	19	AI379201	AI379201 td01f09.x
40	429	15.3	478	159	N28029	N28029 yx50h11.r1
41	426.4	15.2	428	24	AI761741	AI761741 wg67g05.x
42	426.4	15.2	428	24	AI766228	AI766228 wh68c11.x
43	418.2	14.9	428	4	AA281348	AA281348 zs94g12.r1
44	418	14.9	433	159	N35179	N35179 yx83e12.r1
45	415.8	14.8	430	11	AA768829	AA768829 nz64g08.s

ALIGNMENTS

```
CDS             <1..2168
                /note="putative"
                /codon_start=3
                /protein_id="BAB25835.1"
                /db_xref="GI:12843038"
                /translation="GVPFQGEDRNFSPOIKVNLNHRKGLGPSQODPNRFDRDRLFSVV
                SRGVPEELTGLLEYLRRTSKYLTLDSAYTEGSTGKTCLMKAVLNLQDGVNACILPLLQI
                DRDSGNPOPLVNAQCTDEFYHGHSALHIAIEKRSLWCVKLLVENGANVHIRACGRFFQ
                KHQGTCFYFGELPLSLAACRKQWDVVTYLLENPHOPASLEATDSLGNTVLHALVMIAD
                NSPENSALVIHMDSLLQMGARLCPTVQLEDICNHOGLTPLKLAAREGKIEIFRHILQ
                REFSGLYQPLSRKFTEWCYGPVRVSLYDLSSVDSWEKNSVLEIIAFHCKSPHRHRMVV
                LEPLNKLLQEKWDRLIPRFFENFACYLVYMIFFTVAYHQPSLEQPAIPSSKATFGDS
                FVETEWYLLILILGGIYLLGLMYFWRRLFIWISFMDSYFEILFLVLVYILVSQVLR
                FGFAVLVSLSREARSPKAPEDSNTTVTERPTLGQEEEPVPYGGILDASLELFKFTIG
                MGELAFQEOLRFRGVVLLLLAYVLLLTYVLLLNMILALMSTTVNSVATDSWSIWKLQK
                AISVLEMENGYWWCRRKRHRAGRLKVGTKGDGIPDERWCFRVEEVNWAAWEKTLPTL
                SEDPSGAGITGYKKNPTSKPGKNSASEEDHLPLQVLQSH"
BASE COUNT      497 a        630 c        612 g        531 t
ORIGIN

Query Match     53.6%;  Score 1505.8;  DB 192;  Length 2270;
Best Local Similarity   81.1%;
Matches 1822;  Conservative   0;  Mismatches  397;  Indels   28;  Gaps   5;

Qy    487  ttccagggcccaggaccggaaattcgcccctccagataagagtcaacctcaactaccgaaag  546
           |||  | |||| |||||||| ||| ||||||||| ||||||||||||||||||||| |||||
Db     12  TTCCAGGGAGGAGAGACCGGAACTTCTCCCCTCAGATTAAAAGTGAATCTCAACACCGAAAG   71

Qy    547  ggaacagttgccagtcagccggaatcacaaccgatttgaccagaatcggctcttcaatgcg  606
           |||   ||| ||||||||||||||||||| |||| |||||||||||||  ||| ||||| |
Db     72  GGACTCGGTGCCAGTCAGCCAACAGGACCCAAATCGGTTCAACAGTCTTCAGTGTG  131

Qy    607  gtctcccgggggtgtcccaggaatctggctggaagtacctggacaagaacagc  666
           |||||| ||||| ||| ||||| |||||||||||| ||||| |||||||||||
Db    132  GTCTCCCGGGGGTGTCCCGGAGGAGCTGACTGGACTTAGAGTACCTGCGCCCAGC  191

Qy    667  aagtacctcaccgactccggaatacacaagagggggctcaacagtaagagacgtaagagac  726
           |||||||||| |||||||||||| || ||||  |||| |||||||| |||||| ||||| |
Db    192  AAGTACCTCACTGACTCCGGCATACACACGAGAAGGCTCAACAGTGCTGCCTGATGAAG  251

Qy    727  gctgtgctgaacttaaggacggagtcaaatgcctgcccactgctgcagatgac  786
           |||||||||||| ||||| || |||| |||| ||| ||||||||| |||||||
Db    252  GCTGTGCTGAACCTTCAGGATGGGGTCAAATCCGTGTATCCCGTGCTCCAGATTGAC  311

Qy    787  aggggactctggcaaatcctcagccccctggtaaatgcccagtgcacagagtgactattaccga  846
           ||||  |||||||| |||||||||||| ||| ||| |||| ||||| ||||||| |||||||
Db    312  AGGGATTCCGGCAATCCTCAGCCCCCTTGTCAATGCCCCAGTGCGCACCGATGAGTTCTACCGA  371

Qy    847  ggccacagcgtctgcacatccccatgagaagagaggagctcaaacacccctg  906
           ||||||||| ||| |||| | ||||||| |||| ||| |||  |||||||
Db    372  GGCCACAGTGCCCTGCACATCCCATAGAGAAGAGATCCAAAAGGTTCTGGTCGAAGCTCGTG  431

Qy    907  gtgggagaatggggccaaatgtgcaatgtgccctggttcggagggggccaa  966
           |||||| ||||||||||||| ||  |||||||||| |||| ||||||| |||
Db    432  GTAGAGAAATGGGGACAGTTCACATCCGAGCCCTTTCTTCTTCCAGAAGGGCCGA  491

Qy    967  ggggacttgcttttattcgggttgagctcaacccatgga  1026
           || |||||| ||| |||||||| ||||||| ||||||||
Db    492  GGAACTTGTTTCTATTTGGAGCTACCTCTTTCTGGAGAGCCAAGCAGTGG  551

Qy   1027  gatgtggttaagctactcctcctggagaaccccagcctgcaggccactgac  1086
           ||| ||||||  ||| ||||||||||||| |||||||||| |||| |||||
Db    552  GATGTGGTGACCTACCTCCTGGAGAACCCCCAGCCTGGAGCCCACCGAC  611

Qy   1087  tcccagggccaacaagtcctgcatctcggacaaactcagctgagaac  1146
           ||||||| ||||||||| ||||||||||||| ||||||||||||| ||
Db    612  TCCCTGGGGCAACACAGTCCTGCCATGCTCTTCGTTAATGAGACAACTCACC  671

Qy   1147  attgcactggtgaccagccatcatgatggggctcctccaaggcccgcctctgccct  1206
           || |||| ||||||| ||| |||||||||||||| ||| |||||||| |||||||
Db    672  AGTGCCCTGGTGATCCAACATGTATGACAGCGTTCCCAAATGGGGGCCGGCCTCTGCCC  731

Qy   1207  accgtgcagcttgaggacatccgcaacctgcaagatgtcacgctctgaagctggccgcc  1266
           || | |||| ||||||||||| |||||||||||| ||  ||||||| |||||||| |||
Db    732  ACTGTACAGCTTGAGGATATCCGCAACCTGCAAGATGTCCCACACCCTGAAAGTTGGGTCGCC  791

Qy   1267  aaggaggggcaagatcagaagatttcaaggcacatcctgcaggcgcgggagtttcaggactgagc  1326
           ||| |||||| ||| ||| || ||| |||||||| ||||| |||||| ||| |||||||
Db    792  AAGGAAGGTAAAAATTGAGAATCTTCAGGCACATCTCAAGGGCCGTTCAGGGGCCATTCC  851

Qy   1327  ca---cctttccccgaaagttcaccgagtctgctctgtatggggctcgtctgtat  1383
           || |  | ||||||||| ||||  |||||||||| ||| |||||||| ||||||
Db    852  CAGCCCCTTTCCCGAAAGTTCACCGAGTCTGCTATGGGGCTCGTGTTGATCCTGCTT  911

Qy   1384  gacctggctcctgtgacagctgtatgggccgtgtcccggagtcgtcactgtat  1443
           || |||||| | ||| | |||  || ||  |||||||| ||| | |||||  |
Db    912  GACCTGTCCTCTGTGGACAGTTGGGAAAAGAACTCGGTGCTCGAAATCATCGCTTTTCAT  971

Qy   1444  tgcaagagccccgcaccgaatcgtcttttggagccccctgaacaaactgctcag  1503
           |||| |||||| || || || |||| |||||||| ||||| |  |||| ||||
Db    972  TGCAAGAGCCCGCACCGCACCCGCATGGTGGTTTTGAGCCACTGAACAAGCTTCTGCAG  1031

Qy   1504  gcgaaatgggatctgaatcctgcccaagttcctcttaaacttcctgtgtaatctgatctac  1563
           || |||||||||| |||||| ||||||||||| |||| |||| ||||| ||||| |||
Db   1032  GAGAAACTGGGATCGGGTCCAATCCCAAGATTCTTCTTCAACTTCCGGTTCTGGTCTAC  1091

Qy   1564  atgttcatcttcaccgctgttgcctaccatcagcctaccctgaagagcaggccgccccct  1623
           ||||||||||||||| | |||||||| || |||||| || ||||||||| |||| |||
Db   1092  ATGATCATCTTCACACCAGTTGCCTACCACCAGCCTTCCCTGGAGCAGCCAGCCATTCC  1151

Qy   1624  cacctgaaagcggagttggaaaactccatgtctgctgacagggccaacatccttatcctgcta  1683
           ||| |||| ||| |||||||| || ||||| |||||| | |||  ||||| |||||| ||
Db   1152  TCATCAAAAAGCGACTTTTGGGGACTCCATGAGTGTGGGCCACATTTCTGCTTTTCCT  1211

Qy   1684  ggggggatctacctcctcctcgggggccgtgctgtggggcgcacagtgttcatc  1743
           ||||| ||| ||||||||||| ||||| ||| |||||| | |||||| |||||
Db   1212  GGGGGTATTTACCTCCTTCACTGGGGCCAGCTGGTACTTTTTGGGGCGCGCCTGTGTTCAATC  1271

Qy   1744  tggatctcgttcatagacagctacttttgaaaatcctcttcctgttccaggcaca  1803
           ||||| |||||| |||||| ||||| ||| |||||| ||||| ||||| || |||
Db   1272  TGGATCTCGTTCAATGGACAGTTACTTTGAAAATCCTCTTCCTGTTCCAGGCGCGGCTGCTCACA  1331

Qy   1804  gtggtgtcccaggtgctgtgtttcctggccacctgagtggctggtggttctt  1863
           ||||||| ||||| ||||||| || ||||||| ||||| |||||||| ||||
Db   1332  GTGGTGTCCCAGGTGCTGCGCTTCCGTGGAACTGAAGTGGTACCTCCCCGTTAGTGTCA  1391

Qy   1864  gcggctggttgctgggctgaacctgctttactatacgtggctccagcacacaggc  1923
           || |||||||||||||||  |||| |||||||||||| ||||| ||||||| |||
Db   1392  TCCCTAGTGGTGGCTGGCTGGGCTTGAACCTGCTTTATTATACACGTGGCTTCAGCAGACAGGC  1451

Qy   1924  actacagtgtcatgaccagagaggtcatcctggcggaacctgctcgctcttctgatc  1983
           ||||| || ||| || ||||| ||||| ||||| |  ||||| ||  ||| ||||
Db   1452  ATCTACAGTGTCATGACCAGAGAGGTCATCCTGGCGGAAAAGGTCTCTGGTC  1511

Qy   1984  tacttagtctccttctttcggtcttccgctgtagctcgtggtgagctgcaggaaggcttgg  2043
           |||||||| || ||||||| || |||||| || |||| ||| ||||| || |||||||||
Db   1512  TACCTAGTCTCTTCCTTTCGGTCTTCGCTGTAGCTCGTGGAGAGCTGCAGGAAGGCTTTGA  1571

Qy   2044  cgccccgaagctcctacagaggccccaatgccacagagtcaagtgcagccaatgagggacag  2103
           || || ||||| |||||||||||||||| || |||||| |||| ||| |||  ||||||||
Db   1572  AGTCCCAAAGCCCCTGAAGATAGCCAACACCACTGTGACGGAAAGACCCACGCTGGGGCAG  1631

Qy   2104  gaggacgagggcaacgggccagtacagggttatcctggaagcctctc  2163
           ||||| ||| ||||||   |||||| ||||| |||| ||||| ||||
Db   1632  GAGGAGGAG-----CCAGTCCCATATGGGGGCATTGGGATCGCCCTAGAGCTGTTC  1685

Qy   2164  aaattcaccatacggcatgggtcgagctggcctccaggaggcagctgcacttccgcgcggcatg  2223
           ||||||||||| |||||||| |||||||| ||| |||||||||| ||||| ||||||
Db   1686  AAGTTCACCATTGGTATGGGTGAGCTGGCCTTTCAGGAGCAACTGCGTTTTCGGGGTG  1745

Qy   2224  gtggctgctgctgctgggcctaacgtgctgctcacctacatcctgctcaacatgtc  2283
           ||||||||| |||||||||  |||||||||||||| ||| |||||||||||||||
Db   1746  GTGCTGCTGCTGCTGGGCTTACGTGCTGCTCACCTACGTCCTCCTGCTCAACATGTC  1805
```

```
RESULT 2

LOCUS       AU141855      793 bp      mRNA                        25-OCT-2000
DEFINITION  AU141855 THYRO1 Homo sapiens cDNA clone THYRO1001346 5', mRNA
            sequence.
ACCESSION   AU141855
VERSION     AU141855.1  GI:11003376
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 793)
  AUTHORS   Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y.,
            Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and
            Isogai,T.
  TITLE     HRI human cDNA project
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Takao Isogai
            Genomics Laboratory
            Helix Research Institute
            1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
            Tel: 81-438-52-3951
            Fax: 81-438-52-3952
            Email: genomics@hri.co.jp
            HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
            Research Institute; cDNA library construction: Department of
            Virology, Institute of Medical Science, University of Tokyo, and
            Helix Research Institute.
FEATURES             Location/Qualifiers
     source          1..793
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="THYRO1001346"
                     /clone_lib="THYRO1"
                     /tissue_type="thyroid gland"
                     /note="Vector: pME18SFL3"
BASE COUNT      184 a     229 c     230 g     147 t     3 others
ORIGIN
```

```
Qy   2284  atcgccctcatgagcgagccagtcaacagtgtcgccactgacagctgggactctggaag  2343
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1806  ATTGCCCTCATGAGCTGAGACTGTCAAACACGTTGCCACTGACACGCTGGGAGCATCTGGAAG  1865

Qy   2344  ctgcagaaagccaccatctctgtcctggagatggagaatggctattggtggtggcacgaagaa-  2402
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db   1866  TTGCAGAAAGCCACCATCTCTGTCCTTGGAGATGGAGAATGGTTACTGGTGGTCCAGGAGAAA  1925

Qy   2403  --gcagcgggcaggtgtgtgatgctgcactaagccagatgggcagcccggatgag  2460
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1926  AGGCACATCCGCCAGGGAGAGGCTGCCTGAAAGTTGGCACCAAAAGGGGGTAAACCTGATGAG  1985

Qy   2461  cgctggttgcttcaggtgtgaggtgaactgggctcttcatgatgggagagagctgcctacg  2520
           |||||||| || |||||||||||||||||||||||||||||| |||||||||||
Db   1986  CGCTGGTTGCTTCAGGTGTGGAGGTGGAGGAAGTAAACTGGGCTCCATGGGGAGAAGACCCTTCCCACC  2045

Qy   2521  ctgtgtgaggacccgtcaggggtgtccctgcgaacctctcgagaaacccctgtcctggct  2580
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2046  TTATCTGAGGATCCATCAGGGGTGGCAGGCAGCCATCAGGGGAGGAGGTGTTATAAAAGAAACC  2094

Qy   2581  tcccctccaaggaggatggaggaggatggtgcctctgaggagaaaactatgtgccgtccagctc  2640
           |||||| |||||||||||||||||||||||||||||| ||| |||||
Db   2095  -CAACCTCTAAACCTGGGAAAGAAACAGTGCCTCAGGAAGACAACCATCTGCCTCTTCAGGTC  2153

Qy   2641  ctccagtccaactgatggcccagagtgcaacagga----ggccaagaggacaagaagagga  2696
           ||||||||||||||||||||||||||||||||         || ||| ||| ||| |||
Db   2154  CTCCAGTCCCACTGACGGTCCCAGATGCGGCGCAGGCCAGAAGCCTGGGAGTAGAGTAGGGAA  2213

Qy   2697  tcttccaaccacacatctgctggctctg  2723
           ||| ||| ||||||||||||| |||||
Db   2214  TTTTGCCACGCCACACCCGAGGCTACTG  2240
```

```
RESULT 3

LOCUS       BG254423      953 bp      mRNA                        13-FEB-2001
DEFINITION  6023690921F1 NIH_MGC_91 Homo sapiens cDNA clone IMAGE:4477157 5',
            mRNA sequence.
ACCESSION   BG254423
VERSION     BG254423.1  GI:12764239
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
Query Match  26.9%;  Score 755.8;  DB 108;  Length 793;
Best Local Similarity  98.5%;  Pred. No. 3.5e-161;
Matches   782;  Conservative   0;  Mismatches  10;  Indels   2;  Gaps   2;

Qy   203  gagaagctccgggactcccagcccagccgcccacgcccctggcctcaagcctgcgggctcaagtc  262
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1  GAGAAGCTCCGGGACTCCCAGCCCAGCCGCCCACGCCCCTGGCCTCGGGGCTCCAGTC  60

Qy   263  aggccaacaccgacgcgcaagctgggaggaaaacacaaggaaaccctttgacatctcaatctgcaca  322
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61  AGGCCAACACCGACGCGCGCAAGCTGCTGGGAGGAAGACAAGGAGACACCTTGACACATCCATCTGCCACA  120

Qy   323  gaggtcctggctggaccggacagcagcctctcctcctaggagtgactcaccctcacccaccctc  382
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121  GAGGTCCTGGCTGGACCGGACTTCCACACTCCTAGGATGACTCACCCTCACCCTC  180

Qy   383  cagtcttcaggtgttgggagacattagatggaggccaagaagaagatggctctgagtcgggacagag  442
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181  CAGTTTTCAGGTTGTGGACACATTAGATGGAGGCCAAGAAGAAGAATGGCTCTGAGGGGGACAGAG  240

Qy   443  gaaagctggattttgggagcgggggctgccctcccatgaagtcacacagttccaggggaagacc  502
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241  GAAAGCTGGATTTTGGGAGCGGGGGCTGCCCTCCCATGGAGTCACAGGTTCCAGGGGACGAGAG  300

Qy   503  ggaaattccgcccctcagaataagaagtcaaacctcaacctaccgaaagagaacaggtgccagtc  562
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301  GGAAATTCCGCCCCTCAGATAACAGATCAAACCTCAAACTACCGAAAGGGAACAGGTGCCAGTC  360

Qy   563  agccggatccaaaaccgatttgaccgaactcgggctcttcaaatgcggtctcccgggggtgtcc  622
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361  AGCCGGATCCAAAACCGATTTGACCGAACTCGGCGTCTTCAATGCGGTCTCCCGGGGGTGTCC  420

Qy   623  ccgaggagatctggctggactctcaagatctggagcaagcaagagctacctcacggact  682
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421  CCGAGGATCTGGCTGGACTTCCACACTCCTAGGATGACTCACCCTCACCCCGGACT  480

Qy   683  cggaatacacagaagggctccacagttaagaacgtgcctgatgaaggctgtgtgctgaactta  742
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481  CGGAATACACAGAAGGGGCTCCACAGGTAAGACCTGCCTGACATGAAGGCTGTGTGAACCTTA  540

Qy   743  aggacacggagtcaatgcctgcattctgccactgctgcagatcggacaaggggactctggcaatc  802
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541  AGGACGCGGAGTCAAATGCCTGCATTCTGCCACTGCTGCAGATCGGACAGGGGACTCTGGCAATC  600

Qy   803  ctcaagcccctggtaaatgcccagcagatgactaccacggagcccagcgccctgc  862
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601  CTCAGCCCCTGGTAAATGCCCACTGACACATGCTATTACCCGAGCAACAGCGCTGC  660

Qy   863  acatcgccaatgagaagaaggagtcctcgagtctgtgtgaagctcctggtggagaatgggggcca  922
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661  ACATCGCCCATTGAGAAGAAGGAGTCTCGAAATGTGTGAAGCTTCTGGTGGAAGAATGGGGGCCA  720

Qy   923  atgtgcatgcccggagccgcgcctcttcaggaagggccaaggggacttgcttttatt  982
          ||| |||||||||||||||||||||||||| || |||||| | ||
Db   721  ATGTGCATGCCC-GGCCTGCGCGCCG-TTNTTCAGAANNGGNCAAGGGGACTTGCTTTNAAT  778

Qy   983  tcggtgaagctaccc  996
          ||||||||||||||
Db   779  TCGGGAAGCTACCC  792
```

```
REFERENCE   1  (bases 1 to 953)
  AUTHORS    NIH-MGC http://mgc.nci.nih.gov/.
  TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL    Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10305 row: j column: 06
            High quality sequence stop: 676.
FEATURES             Location/Qualifiers
     source          1..953
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4477157"
                     /clone_lib="NIH_MGC_91"
                     /tissue_type="adenocarcinoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: prostate; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                     Average insert size 1.4 kb. Library enriched for
                     full-length clones and constructed by Life Technologies.
                     Note: this is a NIH_MGC Library."
BASE COUNT    201 a    284 c    269 g    199 t
ORIGIN

Query Match    23.28;    Score 651.4;  DB 175;  Length 953;
Best Local Similarity    95.8%;    Pred. No. 1.9e-137;
Matches  702;  Conservative  0;  Mismatches  26;  Indels  5;  Gaps  3;

Qy   852 cagcgctctgcacatcgccaattgagagaagagagtctgca-gtgtgtgaagctcctggtgg  910
         |||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db     1 cagcgctctgcacatcgccaattgagagaagagagtctgcag-gtgtgtgaagctcctggtgg  60

Qy   911 agaatgggcaacaagtgcatccccgggcccgggccgcttcttcccagggaaggcccaaggga  970
         |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db    61 agaatgggcaacaagtgcatccccgggcccgggcctcttcttcccagaaggcccaacaagga  120

Qy   971 cttgctttttatttcggtgagctaccctctcttgggcgcttgcaccaagcagtgggatg  1030
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 cttgctttttatttcggtgagctaccctctcttgttgccgcgcttgcaccaagcagtgggatg  180

Qy  1031 tggttaagctacctcctggagaaccacacagccccgcccagccctgcagccactgacctccc  1090
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 tggttaagctacctcctggagaaccacacagccccgcccagccctgcaggccactgactccc  240

Qy  1091 agggcaacaagtcctgcatccccatagtgatgtctcggacaactcaagctgagaaacattg  1150
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 agggcaacaacatcctccatgccctagtgatctggacaactcagctgaacattg  300

Qy  1151 cactggtggaccaacatgtatgtggctcctccaagctgggccgacctctcccctaccg  1210
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 cactgtgaccagcagtgtatgggctcctccaagctgggccccctgcccctaccg  360

Qy  1211 tgcagcttgaggacaatccgcaacctgcaggatctcacgcctctgaagctggccgcccaagg  1270
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 tgcagcttgaggacaatccgcaacctcgaggatctcacgcctctgaagctggccgcaagg  420

Qy  1271 agggcaagatcgaggatgtttcaggcacacatcctgcacgcgggagtttcaggactgagccacc  1330
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 agggcaagatcgagatttcaggcacacatcctcaggactgagtttcaggactgagccacc  480

Qy  1331 tttcccgaaagttcaccgagtggtgtgctaatggcctgtctccggtgtcgtatgacctg  1390
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 tttccgaaagttcaccgagtgtgtgctaatgggcctgtctccggtgtcgtatgacctg  540
```

```
Qy  1391 cttctgttggacaagctgtcgaggagaactcagtgctgctggagatcattgccttcattgcaaga  1450
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 cttctgttggaccggctgtgtggaggagaactcagtgctggagaactcattgccttcattgcaaga  600

Qy  1451 gcccgcaccgaccaccgaatggtcgtttgggagccccctgaacaaactgctgctcaggcgaaat  1510
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 gcccgcaccgaccaccgaatggtcgtt-tggagccccctgaacaaactgctgctcaggcgaaaa  659

Qy  1511 gggatctggtcatcccccaagttcttcttcaaacttcctgtgtaatctgtatcacatgttca  1570
         |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db   660 tggggatctggtcatcccccaagttcttcttcaaacttcctgtgt---anctgavctacantgttct  716

Qy  1571 tcttcaccgctgt  1583
         |||||||||||||
Db   717 CCTTCACCGCTGT  729
```

```
RESULT  4
BF689928
LOCUS       BF689928        840 bp    mRNA            EST       22-DEC-2000
DEFINITION  601B6482F1 NIH_MGC_49 Homo sapiens cDNA clone IMAGE:4298445 5',
            mRNA sequence.
ACCESSION   BF689928
VERSION     BF689928.1  GI:11975336
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 840)
  AUTHORS    NIH-MGC http://mgc.nci.nih.gov/.
  TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL    Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC/DCTD/DTP
            cDNA Library Preparation: Ling Hong/Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1154 row: c column: 22
            High quality sequence stop: 662.
FEATURES             Location/Qualifiers
     source          1..840
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4298445"
                     /clone_lib="NIH_MGC_49"
                     /tissue_type="melanotic melanoma, high MDR (cell line)"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: skin; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming. Directionally cloned
                     into EcoRI/XhoI sites using the following 5' adaptor:
                     GGCACGAG(G). Size-selected >500bp for average insert size
                     1.8kb. Library constructed by Ling Hong in the laboratory
                     of Gerald M. Rubin (University of California, Berkeley)
                     using ZAP-cDNA synthesis kit (Stratagene) and Superscript
                     II RT (Life Technologies). Note: this is a NIH_MGC
                     Library."
BASE COUNT    174 a    256 c    258 g    152 t
ORIGIN

Query Match    23.1%;    Score 647.8;  DB 168;  Length 840;
Best Local Similarity    96.4%;    Pred. No. 1.2e-136;
Matches  716;  Conservative  0;  Mismatches  22;  Indels  5;  Gaps  5;

Qy    21 ggagagttaagctcccgttcctccaccgtgccgcagtggccaggtggctgaggtgaccga  80
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 ggagagttaagctcccgttcctccaccgtgccgcagtggcctggcctgaggtgaccga  60
```

```
Qy    81 gagaccagaacctgcttgcttggagcttagtgttctaagagtctggggaggggaggttccgcgc 140
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 GAGACCAGAACCTGCTTGCTTGGAGCTTAGTGTTCTAAGAGTCTGGGGAGGGGAGGTTCCGCGC 120

Qy   141 tcctctgtcagcggccgcagcccctccgggcttcacttcctcccgcgcagccccctgcta 200
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 TCCTCTGCTCAGCGCGCCGCAGCCCCTCCGGGCTTCACTTCCTCCCGCGCAGCCCCTGCTA 180

Qy   201 ctgagaagctccgggatcccagacgcgcacgcccacccctggcctcagccgtcgggggtccaag 260
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 CTGAGAAGCTCCGGGATCCCAGCAGCGCGCACGCCCACCCCTGGCCTCAGCCGTCGGGGGTCCAG 240

Qy   261 tcaggccaacaccagcgcgcagctgggaggaagacccttgacactctgca 320
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 TCAGGCCAACACCAGCGCGCAGCTGGGAGGAAGACAGGACCCTTGACACTCTGCCATCTGCA 300

Qy   321 cagaggtcctggctggaccgagcagcagccctcctcctcctcaggtgatgtgacctcaccctcagctc 380
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 CAGAGGTCCTGGCTGGACCGAGCAGCAGCCCTTCCTCCTCCTAGGATGACCTCACCCTCAGCTC 360

Qy   381 tccagtttcaagttggagacattagatggagggccaagaagaagatggtctgaggcggacag 440
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 TCCAGTTTTCAAGTTGGAGACATTAGATGGAGGGCCAAGAAGAAGATGGCTCTGAGGCGGACACAG 420

Qy   441 aggaaagctggattttgggacgcgggctgcctcccatggagtcacagttccagggggagga 500
         ||||||||||||||||| |||||||||||||||||||||||||||| |||||||||||||
Db   421 AGGAAAGCTGGA-TTTGGGACGCGGGCTGCCTCCCATGGAGTCACAGTTCCA-GGGGAGGA 478

Qy   501 ccggaaattcgcccctcagataagagtcaacctacccaaaggaacaaggttgccag 560
         |||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db   479 CCGGAAATTCGCCCCTCAGATAAGAGTCAA-CTCAACTACCCGAAAGGAACAAGGTGCCAG 537

Qy   561 tcagccggatcc-aaaccgatttgaccgagatcggtcttcaatgcggtctcccgggtg 619
         |||||||||||| |||||||||||||||||||||||||||||||||| ||||||||||
Db   538 TCAGCCGGATCCAAAAACCGATTTGACCGAGATCGGTCTTCAATGCGGTCTCACCCGGGTG 597

Qy   620 tccccgaggagtctggcttccagagtacctgagcaagcaagtacctcaccg 679
         ||||||||||||||||||||||||||||||||||||||||||||||| ||
Db   598 TCCCCGAGGAGTCTGGCTGGACTTCCAGAGTACCTGAGCAAGCAAGTACCTCA-CG 656

Qy   680 actcggaatacacagaggggctccacacaggtaagacgtgctcctgatgaaggctgtgtctgaacc 739
         |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db   657 ACTCGGAATACACAGAGGGGCGCCACAGGTAAGACGTGCTCCTGGATGAAGCTGTGTGCTGACCT 716

Qy   740 ttaaggacggagtcaatgcctgc 762
         |||| |||||||||||| ||||
Db   717 TAAGGAGGAGGTCATCCTGCGTTC 739
```

RESULT 5
LOCUS   AW976725   686 bp   mRNA   EST   02-JUN-2000
DEFINITION EST388834 MAGE resequences, MAGO Homo sapiens cDNA, mRNA sequence.
ACCESSION   AW976725
VERSION   AW976725.1 GI:8167958
KEYWORDS   EST.
SOURCE   human.
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 686)
  AUTHORS   Hegde,P., Qi,R., Abernathy,K., Dharap,S., Gaspard,R., Gay,C., Holt
            ,I.E., Saeed,A.I., Sharov,V., Lee,N.H., Yeatman,T.J. and
            Quackenbush,J.
  TITLE   Assessment of gene expression patterns in a model of colon tumor
            metastasis using a 19,200 element cDNA microarray
  JOURNAL   Unpublished (2000)
COMMENT   Contact: John Quackenbush
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 3528
            Fax: 301 838 0208
            Email: johnq@tigr.org

FEATURES   Location/Qualifiers
     source   1..686
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="MAGE resequences, MAGO"
                     /note="Vector: pBluescriptSKm"
BASE COUNT   124 a   195 c   199 g   168 t
ORIGIN

Query Match   22.4%; Score 629.8; DB 123; Length 686;
Best Local Similarity   96.6%;   Pred. No. 1.5e-132;
Matches 654; Conservative   0; Mismatches   22; Indels   1; Gaps   1;

Plate: 364
Seq primer: Forward.

```
Qy   272 ccgacgcgcagctgggaggaagacaggaccccttgacatctgcacacagggtcctg 331
         |||| ||| ||||||||||||||||||||||||||||| -AAATCTCCCATCGCCAAAGGTCCTG
Db   682 CCACCGCCCAAACTGGGAGGAAAACAGGACCCTT-AAATCTCCCATCGCCAAAGGTCCTG 624

Qy   332 gctggaccgagcagcctcctctcctctcctaggatgacctcaccctccagctctccagtttca 391
         ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   623 GCTGGCCGACCAACCAACCTCTCCTCTCCTAGGATGCACCTCACCCTCCAGCTCTCCAGTTTCCA 564

Qy   392 ggttgggagacattagatggaggggccaagaagaagatggctctgaggcggacagaggaaagctgg 451
         ||||||||| |||||||||||||||||||||||||||||||||||||||||||||||| ||
Db   623 GGTTGGGAGACATTAGATGGAGGGCCAAGAAGAAGATGGCTCTGAGGCGGACAGAGGAAAGCTGG 564

Qy   452 atttggggagcgggctgcctcccatggagtcacagttccaggggaggaccggaaattcg 511
         ||||||||||||||||||||||||||||||||||||||||| |||||||||||||||
Db   563 GGTTGGGAGACATTAGATGGAGGGCCAAGAAGAAGATGGCTCTGAGGCGGACAGAGGAAAGCTGG 504

Qy   512 ccctcagataagagtcaacctaacctacccgaaagggaacaaggttgccag 571
         ||||||| ||| ||||||||||| ||||||||||||||||||||||||| ||
Db   503 ATTTTGGGGAGCGGGCTGCCTCCCATGGAGTCACAGTTCCAGGGGAGGACCGGAAATTCG 444

Qy   572 caaaccgatttgaccgagatcggtcttcaatgcggtctcaccaccggggtc 631
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db   443 CCCCTCAGATAAGAGTCAACTCAACTACCCGAAAGGGAACAGGTTGCCAGTCACCCGGATC 384

Qy   632 tggctggacttccagagtacctgagcaagcaagtacctcaccggaataca 691
         ||||||||||||||||||||||||||||||||||||||||||||||
Db   383 CAAACCGGATTTGACCGAGATCGGTCTTCAATGCGGTCTCACCCGGGGTGTCCCCGAGGAT 324

Qy   692 cagagggctccacaggtaagacgtgctcctgatgaaggctgtctgaacctctaaggacgggag 751
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   323 TGGCTGGACTTCCAGAGTACCTGAGCAAGCAAGTACCTCACCGGAATACA 264

Qy   752 tcaatgcctgcatcttccacctgctgactgactgacaggggacagggagactctggcaatcctcagccc 811
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   263 CAGAGGGCTCCACAGGTAAGACGTGCCCAGGTGAAGACCTGGAACCTTAAGGACGGAG 204

Qy   812 tggtaaatgcccagtgcacagatgactattaccgagggccacagccgcctcctgcacatcgca 871
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db   203 TCAATGCCTGCATCTTCCCACTGCCTGCAGATCGACAGGGGACTCTGGCAAATCCTCAGCCCC 144

Qy   872 ttgagaaggaagagtctgcagtgtgtgaagacttcctgtggagaatggggccaatgtgcatg 931
         ||||||||||||||||||||||||||||||||||||||||
Db    83 TTGAGAAGGAGAGTCTGCAGTGTGTGAAGCTCCTGGTCTGGAGAATGGGGCCAATGTGCATG 24

Qy   932 cccgggcctgcggcgc 948
         |||||||||||||||||
Db    23 CCCGGGCCTGCGGCGCGC 7
```

RESULT 6
LOCUS   AU136431   841 bp   mRNA   EST   24-OCT-2000
DEFINITION AU136431 PLACE1 Homo sapiens cDNA clone PLACE1004258 5', mRNA
            sequence.
ACCESSION   AU136431
VERSION   AU136431.1 GI:10996970
KEYWORDS   EST.

```
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE.  1 (bases 1 to 841)
AUTHORS     Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y.,
            Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and
            Isogai,T.
TITLE       HRI human cDNA project
JOURNAL     Unpublished (2000)
COMMENT     Contact: Takao Isogai
            Genomics Laboratory
            Helix Research Institute
            1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
            Tel: 81-438-52-3951
            Fax: 81-438-52-3952
            Email: genomics@hri.co.jp
            This clone has been completely sequenced.
            HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
            Research Institute; cDNA library construction: Department of
            Virology, Institute of Medical Science, University of Tokyo, and
            Helix Research Institute.
FEATURES             Location/Qualifiers
    source           1..841
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="PLACE1004258"
                     /clone_lib="PLACE1"
                     /tissue_type="placenta"
                     /note="Vector: pME18SFL3"
BASE COUNT   193 a    251 c    233 g    151 t    13 others
ORIGIN

Query Match      22.0%;  Score 618.4;  DB 108;  Length 841;
Best Local Similarity   97.9%;  Pred. No. 6e-130;
Matches 644;  Conservative  0;  Mismatches  12;  Indels  2;  Gaps  2;

Qy    32 ctcccgttctccaccgtgcccgctggccaggtgggctgaggtgaccagagaggaccagaac  91
         ||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db     1 ctcccggttctccaccgtgcccgctggccaggtgggctgaggtgaccagagaggaccagaac  60

Qy    92 ctgcttgctggagcttagtgctcaagagctggggaggagttccgcgcgctcctctgctgt  151
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 ctgcttgctggagcttagtgctcaagagctggggaggagttccgcgcgctcctctgcgt  120

Qy   152 cagccgccgcagccccctccccggcgcttcacttcctcccgcgcagcccctgctactgagaggctc  211
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 cagccgccgcagcccctccccggctcacttcctcccgcgcagcccctgctacttgagaagctc  180

Qy   212 cgggatcccagccgccacgccctggcctcagcctgcgggggctccagtcaggccaaca  271
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 cgggatcccagccgccacgccctcggcctcgcggggctccagtcaggccaaca  240

Qy   272 ccgaccgcagctggagggaagcaacgaccgaccctgacatctccaatctgcacagagggtcctg  331
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 ccgaccgcagctgggaggaagacagcggcccctgacatctccatctgcacagaggtcctg  300

Qy   332 gctggaccgagcagcctcctcctccctcaggatgacctcaccctccagctctccagttttca  391
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 gctgggaccgaccagccagctctctctcccctaggatgacctcaccctccagctctccagttttca  360

Qy   392 ggtggagacattagatggaggccaagaagatggctctgaggccgacaagaggaaagctgg  451
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 ggttggagacattagatggaggccaagaagatggctctgaggccgacaagaggaaagctgg  420

Qy   452 atttggaggcggggtgcctcccatgaggagtcacagttcc-aggcgaggaccggaaatc  510
         |||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db   421 atttTggagcgggctgcctcctccatgAGTCacagttccAAAGggcNAGGaCcGgaAATTc  480

Qy   511 gccctcaagataagagtcaacctcaactaccgaaaggaacaggtgccagtcagccggat  570
         |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db   481 GCCCTCAAAATAAGAGTCAACCTCAACTACCGAAAAGGAACAAGTGCCAGTCAAGACCGGAT  540
```

```
RESULT  7
AI801897/c
LOCUS       AI801897       616 bp       mRNA       EST       16-DEC-1999
DEFINITION  tx28h02.x1 NCI_CGAP_Lu24 Homo sapiens cDNA clone IMAGE:2270931 3'
            similar to TR:O35433 O35433 VANILLOID RECEPTOR SUBTYPE 1. ;, mRNA
            sequence.
ACCESSION   AI801897
VERSION     AI801897.1  GI:5367369
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 616)
AUTHORS     NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE       National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
JOURNAL     Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
            cDNA Library Preparation: M. Bento Soares, Ph.D.
            cDNA Library Arrayed by: Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image.html
            Seq primer: -400P from Gibco
            Insert Length: 663    Std Error: 0.00
            High quality sequence stop: 456.
FEATURES             Location/Qualifiers
    source           1..616
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2270931"
                     /clone_lib="NCI_CGAP_Lu24"
                     /tissue_type="carcinoid"
                     /lab_host="DH10B"
                     /note="Organ: lung; Vector: pT7T3D-Pac (Pharmacia) with a
                     modified polylinker; plasmid DNA from the normalized
                     library NCI_CGAP_Lu5 was prepared, and ss circles were
                     made in vitro. Following HAP purification, this DNA was
                     used as tracer in a subtractive hybridization reaction.
                     The driver was PCR-amplified cDNAs from a pool of 5,000
                     clones made from the same library (cloneIDs
                     1414920-1417991 and 1520904-1522439). Subtraction by Bento
                     Soares and M. Fatima Bonaldo."
BASE COUNT   129 a    189 c    169 g    128 t    1 others
ORIGIN

Query Match      21.7%;  Score 610.2;  DB 102;  Length 616;
Best Local Similarity   99.4%;  Pred. No. 4.1e-128;
Matches 612;  Conservative  0;  Mismatches  4;  Indels  0;  Gaps  0;

Qy   571 ccaaaccgatttgaccgagatcggcctctccaatgc-ggtctctcccgggtgtcccgagga  629
         |||||||||||||||||||||||||||||||||||| |||||||||||||||| ||
Db   541 CCAAACCGATTTGACCGAGATNGGCCTCTCCAATGCGGGTCTCTCCCGGGGTGTCCCGAGGA  600

Qy   630 tctggcctggacttccaagtacctgagcaagacaggcaagtacctcaccgaactcggaa  687
         ||||||| |||||||||||||||||||||||| ||||||||||||||||||||||||||| |
Db   601 TNTGGCCTGGACTTCAAGACTANCTGAGCAAGACAGGCAANCAAGGTACCTCACCGACTCGGGA  658
```

```
Qy  2152 ttggagctccttcaaattcaccatccggcatggcggagctggccttccagagagcagctgcac  2211
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   616 TTGGAGCTCTTCAAATTCACCATCCGGCATGGCCGAGCTGGCCTTCCAGAGACAGCTGCAC  557

Qy  2212 ttccgccggcatggtgctgctgctgctggcctacgtgctgctcacctactacctcctgctg  2271
         |||||||||||| |||||||||||||||||||||||||||||||||||||||||||||||
Db   556 TTCCGGCGGCATGGTGCTGCTGCTGCTGGCCTACGTGNTGCTCACCTACTACATCCTGCTG  497
```

```
Qy   2272 ctcaacatgtctcatcgcccctcatgagcgagaccgtcaacagtgtgtcgcccactgacagctgg 2331
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    496 CTCAACATGTCTCATCGCCCTCATGAGCGAGACCGTCAACAGTGTCGCCCACTGACAGCTGG 437

Qy   2332 agcatctggaagctgcagaagcatctctgtcctggagaatgggagaatggctattggtgg 2391
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    436 AGCATCTGGAAGCTGCAGAAGCATCTCTGTCCTGGAGAATGGGAGAATGGCTATTGGTGG 377

Qy   2392 tgcaggaagaagcaacgcgggcagtgtgtgtatgctctgaccgttggccactaagccacagtggcagc 2451
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    376 TGCAGGAAGAAGCAGCGGGCAGTGTGTGTATGCTGACCGTTGGCCACTAAGCCACAGTGGCCGC 317

Qy   2452 ccggatgagcgctggtgcttcagggtggaggtggaggaggtgaactgggccttcatgggagcagacg 2511
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    316 CCGGATGAGCGCTGGTGCTTCAGGGTGGAGGTGGAGGAGGTGAACTGGGCCTTCATGGGAGCAGACG 257

Qy   2512 ctgcctacgctgtgtgaggaccctgtccctcggaacctctcgagaacct 2571
          |||||||||||||||||||||||||||||||||||||||||||||||
Db    256 CTGCCTACGCTGTGTGAGGACCCTGTCCCTCGGAACCTCTCGAGAACCCT 197

Qy   2572 gtcctggcttccctcccctccaaggaggaggatgaggaatggtgctctgaggaaaactatgtgccc 2631
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    196 GTCCTGGCTTCCCTCCCCTCCAAGGAGGAGGATGAGGAATGGTGCTCTGAGGAAAACTATGTGCCC 137

Qy   2632 gtccaagtccctccagtccaaactgatggcccagatgcagcgaggagggacagagca 2691
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    136 GTCCAAGTCCCTCCAGTCCAAACTGATGGCCCAGATGCAGCGAGGAGGGACAGAGCA 77

Qy   2692 gaggatctttccaaccaatctgtgtgggttcccagtgaattcttggtgtggcaaata 2751
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db     76 GAGGATCTTTCCAACCAATCTGTGTGGGTTCCCAGTGAATTCTTGGTGTGGCAAATA 17

Qy   2752 tatatttttcaactaact 2767
          |||||||| |||| |||
Db     16 TATATTATTATCACTAACT 1
```

```
RESULT  8
LOCUS       BF732920        682 bp    mRNA            EST       30-MAR-2001
DEFINITION  nae17b01.x1 NCI_CGAP_Ov18 Homo sapiens cDNA clone IMAGE:3435576 3'
            similar to TR:Q9Y670 Q9Y670 VANILLOID RECEPTOR-LIKE PROTEIN. ;,
            mRNA sequence.
ACCESSION   BF732920
VERSION     BF732920.1  GI:12057995
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 682)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
            R. Emmert-Buck, M.D., Ph.D.
            cDNA Library Preparation: M. Bento
            Soares, Ph.D.
            I.M.A.G.E. cDNA Library Arrayed by: Christa Prange, The
            Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL, send email to:
            info@image.llnl.gov
            Seq primer: -40UP from Gibco
            High quality sequence stop: 412.
FEATURES             Location/Qualifiers
     source          1..682
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3435576"
                     /clone_lib="NCI_CGAP_Ov18"
                     /tissue_type="fibrotheoma"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: ovary; Vector: pT7T3D-Pac (Pharmacia) with a
                     modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGGAGCGGCGGGCGGCGGCGGCGACAATTTTTTTTTTTTTTTT 3'];
                     double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not
                     I and Eco RI sites of the modified pT7T3 vector. Library
                     constructed by Bento Soares and M. Fatima Bonaldo."
BASE COUNT      142 a    208 c    181 g    150 t    1 others
ORIGIN
```

```
Best Local Similarity   94.0%;
Matches    641;  Conservative    0;  Mismatches   39;  Indels   2;  Gaps   2;

Query Match  21.2%;  Score 596.2;  DB 168;  Length 682;
                                            Pred. No. 6.3e-125;

Qy    991 ctacccctcttttggccgcttgcaccaagacagtggggatggtaagctacctcctggag 1050
          ||| | ||| ||| | |||||||||||||||||||||||||||||||||||||||||||||
Db      2 CTAACCGCCCGTGTGCCGCGCGGCGGCGGCGCAAGCACTGGGATGTGGTAAGCTACCTCCTGGAG 61

Qy   1051 aacccacaccagccccgccagccctgcaggccaactgactcccaggggcaacaacagtcctgcat 1110
          ||||||| ||||||||| || |||||||||||||| |||||||||||||||||| || ||| |||
Db     62 AACCCACCACCAGCCCCACCACCCTGCAGGCCACTGACTCCCAGGGGCAACACACCACCCTTCCAT 121

Qy   1111 gccctagtgatgatcctccggagaacaactgcaagctgagaaacattgcaactggtgaacccatgtat 1170
          |||||||||||||| |||||| |||| ||| |||||||||||| |||||||| ||||||| |||||||
Db    122 GCCCTAGTGATGAATCTCGGAGCAAACTCAGCTGAGAACACATTGCACTGGTGACCCAGCATGTAT 181

Qy   1171 gatgggctcctccaagctgtgggcccgcctctgccctaccgtgcagcttggaggactaccgc 1230
          |||||| |||||||||||||||||||| | |||||||||||||||||||||||||| ||||||
Db    182 GATGGGCGCTCCTCCAAGCTGTGGGCCCGCCCTCTGCCCTACCGTGCAGCTTGGAGGACATCCGC 241

Qy   1231 aaacctgcaggatctcacgctctgaagctggccgcccaaggaggggcaaggatcgagagatttc 1290
          ||| |||||||| |||||||| ||||||||| ||||||||||||| |||||||| |||||| ||||
Db    242 AACCTGCAGGGATCTCACGCCTCTGAAGCTGCCGCCCAAGGAGGGCAAGGATCGGAGAGATTTTC 301

Qy   1291 aggcacaatcctgcaggcgggaggtttttcaggggactgagccaacctttccgaaagttcaccgag 1350
          |||||||||||||||||||||| ||| ||||||| |||||||| ||||||| |||||||||||||||
Db    302 AGGCACAATCCTGCAGGCGGGGAGTTTCAGGACTGAGCCACCTTTCCCGAAAGTTCACCGAG 361

Qy   1351 tggtgctcatgggccctgtcccggggtgtcgctgtatgacccggccacgctgtgtga 1410
          |||||| | |||||||| ||||||||||||||||||||||||||||||||||||| ||
Db    362 TGGTGCTATGGGCCCTTCCGGGGTGTCGCTGTATGACCCGGCCACGCTGTGAG 421

Qy   1411 gagaactcagtgctggagatcattgcctctcattgcaagagacccgcaccgaccgaaatg 1470
          ||||||||||||||||||||||||||||| |||||||||||||||||||||| |||| |||
Db    422 GAGAACTCAGTGCTGGAGATCATTGCCTTTCATTGCAAGAGACCCGCACCGACACCGAATG 481

Qy   1471 gtcggtttggagccctgaacacaaactgctgcaggcgaaatggggatctgatcccccaag 1530
          ||||| | ||||||| |||||| |||||||||||||||||||||||||||| ||| | |||
Db    482 GTCGGTGTTGGAGCCCCTGAACAAACTGCTGCAGGCGAAATGGGGATCTGCTCATCCTCAAG 541

Qy   1531 ttcttcttaaacttcctgtgtaatctgatctacacatgttcatcttcaccgtcgttgcctac 1590
          ||||| ||| ||||||||||||||||||||||||||||||||||||||||| | ||| ||||
Db    542 GTTCTCTTTGAACTTCCTGTGTAATCTGATCTACACATGTTCATCTTCACCGGTGGTGCCTAC 601

Qy   1591 catcagcctaccctgaagaagc-aggccgcccctcaacctgaaaaatggggttggaaactc 1649
          ||||||||||| | ||||| ||   |||| ||| ||| ||||||| ||||| ||  | |||
Db    602 CATCAGCCTACTCTTGAGGACCAAGGTCGGTTCCTCACCTGAAAAGCGGAGGGTT-GCAACTT 660

Qy   1650 catgctgctgacggccacatc 1671
          ||||||||||| || | ||||
Db    661 CATGCTGCTGACCGCGTCCATC 682
```

```
RESULT  9
LOCUS       BF940288        588 bp    mRNA            EST       22-JAN-2001
DEFINITION  7o43f06.x1 NCI_CGAP_Kid11 Homo sapiens cDNA clone IMAGE:3577090 3'
            similar to TR:Q9Y5S1 Q9Y5S1 VANILLOID RECEPTOR-LIKE PROTEIN 1. ;,
```

```
ACCESSION    BF940288
VERSION      BF940288.1  GI:12357608
KEYWORDS     EST.
SOURCE       human.
  ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1  (bases 1 to 588)
  AUTHORS    NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE      National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
             Tumor Gene Index
  JOURNAL    Unpublished (1997)
COMMENT      Contact: Robert Strausberg, Ph.D.
             Email: cgapbs-r@mail.nih.gov
             Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
             Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
             found through the I.M.A.G.E. Consortium/LLNL.  send email to:
             info@image.llnl.gov
             Seq primer: -40UP from Gibco
             High quality sequence stop: 492.
FEATURES             Location/Qualifiers
     source          1..588
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3577090"
                     /clone_lib="NCI_CGAP_Kid11"
                     /lab_host="DH10B"
                     /note="Organ: kidney; Vector: pT7T3D-Pac (Pharmacia) with
                     a modified polylinker; Site_1: Not I; Site_2: Eco RI;
                     plasmid DNA from the normalized library NCI_CGAP_Kid3 was
                     prepared, and ss circles were made in vitro. Following HAP
                     purification, this DNA was used as tracer in a subtractive
                     hybridization reaction. The driver was PCR-amplified cDNAs
                     from a pool of 5,000 clones made from the same library
                     (cloneIds 1322376-1323911, 1456007-1456775, and
                     1500552-1502855). Subtraction by Bento Soares and M.
                     Fatima Bonaldo."
BASE COUNT      110 a    173 c    163 g    142 t
ORIGIN

Query Match   20.9%;  Score 588;  DB 171;  Length 588;
Best Local Similarity  100.0%;  Pred. No. 4.4e-123;
Matches  588;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy  361 atgactcacctccagcctctccagcgttcagttcggagacattagatggggaggccaagaa  420
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  588 ATGACCTCACCTCCAGCCTCTCCAGTGTTCAGTTCGGAGACATTAGATGGGGAGGCCAAGAA  529

Qy  421 gatggtctctgaggcggacagagaggaagctgattttgggagccgggctgcctcccatggag  480
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  528 GATGGTCTGAGGCGGACAGAGAGGAAAGCTGGATTTTGGGACGCGGGCTGCCTCCCATGGAG  469

Qy  481 tcacagttccaggcggacggaccggaaattccgcccctcaagataagagagtcaactcaactac  540
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  468 TCACAGTTCCAGGCGGACGGACCGGAAATTCGCCCCTCAGATATAAGAGTCAACTCAACTAC  409

Qy  541 cgaaagggaacaggtgccagtccagccggatccaaaaccgatttgaccggagatcggctcttc  600
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  408 CGAAAGGGAACAGGTGCCAGTCCGGATCCAAAACCGATTTTGACCAGAGATCGGCTCTTC  349

Qy  601 aatgcggtctcccgggggtgtctggactctgggactccagaagtacctgagcaag  660
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  348 AATGCGGTCTCCCGGGGTGTCCCGGAGGATCTGGCTGGGACTTCCAGAGTACCTGGAGCAAG  289

Qy  661 accagcaagtacctcaccggattacacaaggagggctccacaggtaagacgttgcctg  720
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  288 ACCAGCAAGTACCTCACCGGGATTACACAAGGAGGGCTCCACAGGTAAGACGTGCCTG  229
```

```
RESULT 10
LOCUS        BF690154    984 bp    mRNA              EST         22-DEC-2000
DEFINITION   602186482T1 NIH_MGC_49 Homo sapiens cDNA clone IMAGE:4298445 3',
             mRNA sequence.
ACCESSION    BF690154
VERSION      BF690154.1  GI:11975562
KEYWORDS     EST.
SOURCE       human.
  ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1  (bases 1 to 984)
  AUTHORS    NIH-MGC http://mgc.nci.nih.gov/.
  TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL    Unpublished (1999)
COMMENT      Contact: Robert Strausberg, Ph.D.
             Email: cgapbs-r@mail.nih.gov
             Tissue Procurement: ATCC/DCTD/DTP
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
             found through the I.M.A.G.E. Consortium/LLNL at:
             http://image.llnl.gov
             Plate: LLCM1154  row: c  column: 22
             High quality sequence start: 20
             High quality sequence stop: 715.
FEATURES             Location/Qualifiers
     source          1..984
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4298445"
                     /clone_lib="NIH_MGC_49"
                     /tissue_type="melanotic melanoma, high MDR (cell line)"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: skin; Vector: pOTB7; Site_1: Xhol; Site_2:
                     EcoRI; cDNA made by oligo-dT priming. Directionally cloned
                     into EcoRI/Xhol sites using the following 5' adaptor:
                     GGCACGAG(G). Size-selected >500bp for average insert size
                     1.8kb. Library constructed by Ling Hong in the laboratory
                     of Gerald M. Rubin (University of California, Berkeley)
                     using ZAP-cDNA synthesis kit (Stratagene) and Superscript
                     II RT (Life Technologies). Note: this is a NIH_MGC
                     Library."
BASE COUNT      200 a    297 c    284 g    203 t
ORIGIN

Query Match   19.7%;  Score 554;  DB 168;  Length 984;
Best Local Similarity  86.6%;  Pred. No. 2.7e-115;
Matches  752;  Conservative  0;  Mismatches  95;  Indels  21;  Gaps  12;

Qy  1870 gtgctgtggtcttggctgaacctgcttactatacacgtgcttccagcaacacaggcatctac  1929
         |||| ||| ||||||||  ||||| |||||| ||| |||| |||| ||  ||
```

```
RESULT 11
BF43303/c
BF43303

LOCUS       BF433303         543 bp    mRNA            EST       30-MAR-2001
DEFINITION  7g60f10.x1 NCI_CGAP_Lu24 Homo sapiens cDNA clone IMAGE:3702691 3'
            similar to TR:Q9Y5S1 Q9Y5S1 VANILLOID RECEPTOR-LIKE PROTEIN 1. ;,
            mRNA sequence.
ACCESSION   BF433303
VERSION     BF433303.1  GI:11445466
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 543)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
            National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
  TITLE     Tumor Gene Index
  JOURNAL   Unpublished (1997)
  COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
            cDNA Library Preparation: M. Bento Soares, Ph.D.
            cDNA Library Arrayed by: Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL, send email to:
            info@image.llnl.gov
            Seq primer: -40UP from Gibco
            High quality sequence stop: 399.
FEATURES             Location/Qualifiers
     source          1..543
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3702691"
                     /clone_lib="NCI_CGAP_Lu24"
                     /tissue_type="carcinoid"
                     /lab_host="DH10B"
                     /note="Organ: lung; Vector: pT7T3D-Pac (Pharmacia) with a
                     modified polylinker; Plasmid DNA from the normalized
                     library NCI_CGAP_Lu5 was prepared, and ss circles were
                     made in vitro. Following HAP purification, this DNA was
                     used as tracer in a subtractive hybridization reaction.
                     The driver was PCR-amplified cDNAs from a pool of 5,000
                     clones made from the same library (cloneIds
                     1414920-1417991 and 1520904-1522439). Subtraction by Bento
                     Soares and M. Fatima Bonaldo."
BASE COUNT     116 a    164 c    148 g    115 t
ORIGIN

Query Match            19.18;  Score 536.6;  DB 148;  Length 543;
Best Local Similarity  99.38;
Matches  539;  Conservative   0;  Mismatches   4;  Indels   0;  Gaps   0;
```

```
Db   893 GTGCTTGGCTCGGCTGCA-CTTCTTAACTAACCCCTTGGCTTCCAGCACTACAGCCATTGAC 835
Qy  1930 agtgtcatgatcagagaggtcatcctgcgggacctgcgcttcctttctgatctactta 1989

Db   834 ACTGTCAGGATCCAGAAGTCATCCGCGGGAACCCTGTGCGGCGTTCAGTGCTGATCTAGTAG 775
Qy  1990 gtcttcctttctggcttcgctctgtag---ccctggtgagcctgagccaggaggctaggc 2045

Db   774 TCTTCCCTGTTTCGCGGTTCGCTCGCTCGCTTGGAGACCCCTGGTGGAGCCAGGCCAGGAGGTCTTGG 715
Qy  2046 cccggaag----ctcctacaaggcccccaatggcc--acagagtcagtgcagcccatgga-g 2097

Db   714 CGCCCACGAAGCTCCGCTACAGGCCCCCAAAATGCCAACACAGAGTCAGTGCAGCCCATGGAGG 655
Qy  2098 ggacaggaggacgaggggcaacggggcccagatacagaggtgtatcctggaagcctcctgggag 2157

Db   654 GGACAGGAGGAACGAGGGCAACGGGGCCCAGTACAGAGGTGTATCCTGGAAGCCTCCTTGGAG 595
Qy  2158 ctct-tcaaattcaccatcggcatggg-cgagctggcctccaggagcagctgcacttcc 2215

Db   594 CTCTGCTCAAAGTCACCATCACCGGCATGGGTCGAGCTGGCCTCCAGGAGCAGCTGCACTTCC 535
Qy  2216 gcggcatggtggctgctgctg-ctgctggccctacgtgctgctcacctacatcctgctgctc 2274

Db   534 GCGGCCATGGTGGCTGCTGCTCGTGCTGGCCCTACGTGCTGCTCACCTACGTGCTGCTC 475
Qy  2275 aacatgctcatcgccctaatgagcgagaccgtcaacagtgtcgccactgacaagctggagc 2334

Db   474 AACATGCTCATCGCCCTCATGAGCGAGACCGTCAACAGTGTCGCCACTGACAAGCTGGAGC 415
Qy  2335 a-tctggaagctgcagaagccatctctgtccctggagatggagaatggctattggtggtg 2393

Db   414 ATTCTGGAAGCTGCAGAAAGCCATCTCTGTCCCTGGAGATGGAGAATGGCTATTGGTGGTG 355
Qy  2394 caggaagag--cagcggggcaggtgtgatg-ctgaccgttggcactaagccagatggcag 2450

Db   354 CAGGAAGAAGTCACGGGACAGGTGTGATGCTGACCGTTGGCACTAAGCCAGATGGCAG 295
Qy  2451 cccggatgaagcgctggttgcttcagggtgaggaggtgaactggcgttcatgggagcagac 2510

Db   294 CCCCGATGACGGCTGGTGCTTCAGGGTGGAGGAGGTGAACTGGCGTTCATGGGAGCAGAC 235
Qy  2511 gctggcctacgctgtgtgaggaaccgtcaggggcaggtcaggtgtccctggaacctctgag--aacc 2569

Db   234 GCTGCCTACGCTGTGTGAGGAACCGTCAGGGGCAGGTGTCCCTGGAACTCTCGAGAAACC 175
Qy  2570 ctgtccctgcttcccctcccaaggagaggatgaggatgaggatgtgcctctgaggaaaactatgtgc 2629

Db   174 CTGTCCCTGCTTCCCCTCCCAAGGAGAGGATGAGGATGTGCCTCTGAGGAAAACTATGTGC 115
Qy  2630 ccgtccagctcctccaactgatgtgccaggaggccagagagcag 2689

Db   114 CCGTCCAGCTCCTCCAACTGATGGCCCAGATTGCACCAGGAGGCCAGAGGGACAGA 55
Qy  2690 cagaggaatctttccaaccacatctgctg 2717

Db    54 GCAGAGATCAATTCCAACCACATCTGCTG 27
```

```
Qy  2227 ctgctgctgctgtggcctacgtcgtgctgctgctcactacaatcctgctgctcaacatgctcatc 2286
Db   543 CTGCTGCTGTGTTGGCTGGCCTACGTCGTGCTCACTCACTACAATCCTGCTGCTCATC 484

Qy  2287 gccctcatgagcgagaccgtcaacagtgtcgccactgacaagctggagactggaagctg 2346
Db   483 GCCCTCAATGAGCGAGACCGTCAACAGTGTCGCCACTGACAAGCTGGAGCATCTGGAAGCTG 424

Qy  2347 cagaaagccatctctgtccctggagatggagaatggctattggtggtgtcagggaagaagcag 2406
Db   423 CAGAAAGCCATCTCTGTCCCTGGAGATGGAGAATGGCTATTGTGGTGTCCAGCAAGAACCAG 364

Qy  2407 cggcaggtgtgatgctgaccgttggcactaagccagatggcccggatgagcgctgg 2466
Db   363 CGGCAGGTGTGATGCTGACCGTTGGCCACTAAGCCAGATGGCCCGGATGAGCGCTGG 304

Qy  2467 tgcttcagggtgggaggagtgaactggcgttcatgggagcagacgctgcctacgcgtgt 2526
Db   303 TGCTTCAGGGTGGGAGGAGTGAACTGGGCTTCATGGGAGCAGACGCTGCCTACGCTGT 244

Qy  2527 gaggaaccgtcaggggcaggtgtccctggaacctctcgagaaacctgtcctgctccct 2586
Db   243 GAGGAACCGTCAGGGGCAGGTGTCCCTGGAACTCTCGAACACCCTGTCCTGCTCCCT 184

Qy  2587 cccaaggaggatgaggatgtgcctctgaggaaaactatgtgccccgtccagctcctccag 2646
Db   183 CCCAAGGAGGATGAGGATGTGCCTCTGAGGAAAACTATGTGCCCCGTCCAGCTCCTCCAG 124

Qy  2647 tccaactgatggcccagaggaggccagagagacgagagaggatctttccaac 2706
Db   123 TCCAACTGATGGCCCAGATGCACAGCAGGAGGCCAGAGGACAGAGGATCTTTCCAAC 64
```

```
Qy    2707 cacactctgctctgggggtcccaatgaattctggtggcaaatatatatttttcactaac 2766
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      63 CACACTCTGCTCTGGGGTCCCCAGTGAATTCTGGTGGCAAATATATATTTTTCACTAGC 4

Qy    2767 tca 2769
           |||
Db       3 TCA 1
```

RESULT 12

```
LOCUS       BF195711
            BF195711     550 bp    mRNA              EST      03-NOV-2000
DEFINITION  7n87e08.x1 NCI_CGAP_Ov18 Homo sapiens cDNA clone IMAGE:3571671 3'
            similar to TR:Q9Y670 Q9Y670 VANILLOID RECEPTOR-LIKE PROTEIN. ;,
            mRNA sequence.
ACCESSION   BF195711
VERSION     BF195711.1  GI:11082885
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 550)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
            R. Emmert-Buck, M.D., Ph.D. cDNA Library Preparation: M. Bento
            Soares, Ph.D. cDNA Library Arrayed by: Christa Prange, The
            I.M.A.G.E. Consortium DNA Sequencing by: Washington University
            Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL, send email to:
            info@image.llnl.gov
            Trace considered overall poor quality
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1..550
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3571671"
                     /clone_lib="NCI_CGAP_Ov18"
                     /tissue_type="fibrotheoma"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: ovary; Vector: pT7T3D-Pac (Pharmacia) with a
                     modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGGCGGCGGCGCGCGCGCCAATTTTTTTTTTTTTTTT 3'];
                     double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not
                     I and Eco RI sites of the modified pT7T3 vector. Library
                     went through one round of normalization, and was
                     constructed by Bento Soares and M. Fatima Bonaldo."
BASE COUNT  122 a  166 c  149 g  113 t
ORIGIN
```

```
Query Match     19.0%;  Score 532.4;  DB 145;  Length 550;
Best Local Similarity  98.0%;  Pred. No. 1.9e-110;
Matches 539;  Conservative   0;  Mismatches  11;  Indels   0;  Gaps   0;

Qy     984 cggtgagctaccccctcctcttggccgttgcaccaagcagtgggagtgtggtaagctact 1043
           ||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db       1 CGGTGAGCTACCCCCTCCTCTTGGCCCCTTGCACCAAGCAGTGGGATGTGGTAAGCTACCT 60

Qy    1044 cctggagaaccacacaccagcccgcacgcctcgcagcgactcccagcgcaacacagt 1103
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 CCTGGAGAACCCACACCACCAGCCCGCCACCAGCCTGCCAGGGACTCACTCCCAGGGACTACCT 120

Qy    1104 cctgcgatgccctagtgatgatctcggaacaactcagctgagaacaattgcactggtgaccag .1163
```

RESULT 13

```
LOCUS       AW517339/c
            AW517339     527 bp    mRNA              EST      03-MAR-2000
DEFINITION  xp93e02.x1 Soares_NHCe_cervix Homo sapiens cDNA clone IMAGE:2747930
            3' similar to TR:035433 035433 VANILLOID RECEPTOR SUBTYPE 1. ;,
            mRNA sequence.
ACCESSION   AW517339
VERSION     AW517339.1  GI:7155421
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 527)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -40UP from Gibco
            High quality sequence stop: 467.
FEATURES             Location/Qualifiers
     source          1..527
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2747930"
                     /clone_lib="Soares_NHCe_cervix"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: cervix; Vector: pT7T3D-Pac (Pharmacia) with
                     a modified polylinker; Site_1: Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGGCGGCGGCGCGCCTTTTTTTTTTTTTTTTTTTTTT
                     T 3']; double-stranded cDNA was ligated to Eco RI
                     adaptors (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of the modified pT7T3 vector.
                     Library is normalized; constructed by Bento Soares and
                     M.Fatima Bonaldo."
BASE COUNT   96 a  158 c  141 g  131 t  1 others
```

```
Db     121 CCTGCCATGCCCTAGTGATGATCTCGGACAACTCAGCTGAGAACATTGCACTGGTGACCAG 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    1164 catgtataatggggctcctcccaaagttgggggcccccctgccctaaccgtgcaagcttgagga 1223
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 CATGTATGATGGGGCTCCTCCAAGCTGGGGCCCGGGCCCTCCCTACCGTGCAACTTGAGGA 240

Qy    1224 catccgcaacctgcaggatctcacgcccctctgaagctggccgccaagggagggcaagatcga 1283
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 CATCCGCAACCTGCAGGATCTCACGCCCTCTGAAGCTGGCCGCCAAGGAGGGCAAGATCGA 300

Qy    1284 gatttcaggcacatcctgcagcgcgggagtgtttcaggactggccacctttcccgaaagtt 1343
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 GATTTCAGGCACATCCTGCAGCGCGGGAGATTTCAGGACTGGCCACCTTTCCCGAAAGTT 360

Qy    1344 cacggagtgggtgtgtatgggccctctccgggtgtcgctgtatgacctggtcttctgtgacag 1403
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 CACCGAGTGGTGCTGTATGGGCCTGTACGGGTGTCCGCTGTATGACCTGGTCTTCTGTGGACAG 420

Qy    1404 ctgtgaggagaacctcagttgctggagatcattgccttcattgcaagagagcccgcaccgaca 1463
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 CTGTGAGGAGAACTCAGTGCTGGAGGATCATTGCCTTACATTGCAAGAGCCCACACCGACA 480

Qy    1464 ccgaatggcgttttgagccctgaacaaactgctgcaggcgaaatggggatctgctcat 1523
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 CCTAATGGCTGGTTATGGACGCCCCTGAACAAACTGCTGCAGGCGATATGGGATCTGCTCAT 540

Qy    1524 cccaaagtc 1533
           |||||||||
Db     541 GCCAAAGTTC 550
```

ORIGIN

Query Match
Best Local Similarity 99.4%;   Score 522.8; DB 117; Length 527;
Matches 524; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Pred. No. 2.8e-108;

```
Qy  399 gacattagatggaggccaagaagaagatggctctgaggcgacagaggaaagctggatttgg 458
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  527 GACATTAGATGGAGGCCAAGAAGAAGATGGCTCTGAGGCGACAGAGGAAAGCTGGATTTGG 468

Qy  459 gacgggggtgcctcccatggagtcacagttccagggccgaggaccggagaggaaactcgcccctca 518
        |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db  467 GACGGGGGTGCCTCCCATGGAGTCACAGTTCCAGGGCCGAGGACCGGAGANAATTCGCCCCTCA 408

Qy  519 gataagagtcaacctcaactaccgaaaggaacaaggtgccagtcagccggatccaaacg 578
        |||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db  407 GATAAGAGTCAACCTCAACTACCGAAAGGAACAAGGTGCCAGTCAGCCGGATCCAAACCG 348

Qy  579 atttgacacggagatcggcctcttcaatgcggtcccgcggggggtgtccccgaggatctggctgg 638
        ||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  347 ATTTGACCGGAGATCGGCTCTTCAATGCGGTCCCGCGGGGGTGTCCCCGAGGATCTGGCTGG 288

Qy  639 acttccagagtacctgagcaagaccaagcaagtacctcaccgagtcggaatacacagagggg 698
        |||||||||||||||||||||||||||||||||||||||||||||| |||||||||||||||
Db  287 ACTTCCAGAGTACCTGAGCAAGACCAAGCAAGTACCTCACCGACTCGGAATACACAGAGGG 228

Qy  699 ctccacaggtaagacgttgcctgatgaaggctgtgctgaaccttaaggacggagtcaatgc 758
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  227 CTCCAACAGGTAAGACGTTGCCTGATGAAGGCTGTGCTGAACCTTAAGGACGGAGTCAATGC 168

Qy  759 ctgcaattctgccacctgtgcaagatcgaacaatcctcagccccctcggtaaa 818
        |||||||||||||||||| |||| |||||  ||||||||||||||||||||||
Db  167 CTGCAATTCTGCCACTGCTCGCAAATCCTCACCCCCTCGGTAAA 108

Qy  819 tgcccagtgcacagatgactattaccgaggcgctcacactcgccattgagaa 878
        |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  107 TGCCCAGTGCACAGATGACTATTACCGAGGCGCTCACACTCGCCATTGAGAA 48

Qy  879 gaggagtctgcagtgtgtggaagctcctcgttggtgggcaatg 925
        |||||||||||||||||||||||||||||||||||||||||||
Db   47 GAGGAGTCTGCAGTGTGTGGAAGCTCCTCGGTGGTGGGCCAATG 1
```

RESULT 14
BE670695/c
LOCUS   BE670695   536 bp   mRNA   EST   08-SEP-2000
DEFINITION 7e38h05.x1 NCI_CGAP_Lu24 Homo sapiens cDNA clone IMAGE:3284793 3'
            similar to TR:O35433 O35433 VANILLOID RECEPTOR SUBTYPE 1. ;, mRNA
            sequence.
ACCESSION   BE670695
VERSION   BE670695.1 GI:10031236
KEYWORDS   EST.
SOURCE   human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 536)
AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE   National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
JOURNAL   Unpublished (1997)
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
            cDNA Library Preparation: M. Bento Soares, Ph.D.
            cDNA Library Arrayed by: Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL, send email to:
            info@image.llnl.gov
            Seq primer: -40UP from Gibco

FEATURES   Location/Qualifiers
     source   1..536
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3284793"
                     /clone_lib="NCI_CGAP_Lu24"
                     /lab_host="DH10B"
                     /note="Organ: lung; Vector: pT7T3D-Pac (Pharmacia) with a
                     modified polylinker; Plasmid DNA from the normalized
                     library NCI_CGAP_Lu5 was prepared, and ss circles were
                     made in vitro. Following HAP purification, this DNA was
                     used as tracer in a subtractive hybridization reaction.
                     The driver was PCR-amplified cDNAs from a pool of 5,000
                     clones made from the same library (cloneIds
                     1414920-1417991 and 1520904-1522439). Subtraction by Bento
                     Soares and M. Fatima Bonaldo."
BASE COUNT   111 a   161 c   143 g   121 t
ORIGIN

High quality sequence stop: 438.

Query Match
Best Local Similarity 98.5%;   Score 522.2; DB 138; Length 536;
Matches 527; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
Pred. No. 3.9e-108;

```
Qy 2237 tgctggcctacgtgtctgctcaacctacatctgctgctcaacatgctcatcgcccctcatga 2296
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  535 TGCTGCATACGTGCGTCCTCACATACATCATCCCACTGCTCACATAGCTCATCCCCTCATGA 476

Qy 2297 gcgagaaccgtcaaacagtgtcgccactgacagctggagctactgaagctgcagaagcca 2356
        ||||| || ||||| ||||||||||||||||||||||||||||| |||||| |||||||||
Db  475 GCGAGACCCGTCACCAGTGTCGCCACTGACAGCTGGAGCATCTGGAAGCTTCCAGAAAGCCA 416

Qy 2357 tctctgtcctggaagtgggagaatggctattggttgtgtcaggaagaagaagcaggggtg 2416
        |||  ||||||||| ||| |||||||||||||||||||||||||||||||||||||||||
Db  415 TCTATGTCCTGGAGATGGAGAATGGCTATTGGTTGTGTCAGGAAGAAGCAGCAGGCAGGTG 356

Qy 2417 tgatgctgaccgttggcactaagccagatggcagcccgaatggcgttgagcgtgtcttcaggg 2476
        |||||||||||||| |||||||||||||||||||||||| |||||||||||| |||||||||||
Db  355 TGATGCTGACCGTTGCACTAAGCCAGATGGCAGCCCGGATGAGCGCTGGTGCTTCAGGG 296

Qy 2477 tgggaggaggtgaactgggcttcaatgggagcagacagcctcgcctacgctgtgtgaggaccgt 2536
        ||||| || ||||||||||||||| |||||||| ||||||| |||||||||||||||||||||||
Db  295 TGGATGAGAGGTGGAACTGGGCTTCATGGGAGGCAGCAGGACCGGCCCTACGCTGTGTGAGGACCCGT 236

Qy 2537 cagggcaggtgtcctcgaactctcgagaaccctgtcctggcttccctccctcccaaggagg 2596
        |||||||||||||| |||||||||||||||||||||| |||||||||||||||||||||||
Db  235 CAGGGCAGGTGTCCCTCGAACTCTCGAGAACCCTGTTCCTGGCTTCCCTCCCTCCCAAGGAGG 176

Qy 2597 atgaggatggtggctcctgaggagaaactatgtgccgtcagctcctccagtcccaactgat 2656
        |||||||||||||| |||||||||||| ||||||||| ||| |||||||||||||||||||
Db  175 ATGAGGATGGTGGCCTCTGAGGAGAAAACTATGTGCCCGTCCAGCTCCTCCAGTCCCAACTGAT 116

Qy 2657 ggcccagatgcagcaggagggccagaggaggcagaggagaggagtctcttccaacacatctgct 2716
        ||||||||| ||||||||||| |||||||| |||||||| |||||||| ||||||| ||||||||
Db  115 GGCCCAGATCGAGCAGGAGGGACCAGAGGACGAGAGGATCTTTCCAACCACCATCTGCT 56

Qy 2717 ggctctggggtcccagtgaattctggttggcaaatataatatattttcactaactcaaa 2771
        |||||||||||||||||||||||||||| ||||||||| ||||||||||||||||||||
Db   55 GGCTCTGGGGTCCCAGTGAATTCTGGTGGGCAAATATATATATTTTCACTAACTCAAA 1
```

RESULT 15
BF436096/c
LOCUS   BF436096   513 bp   mRNA   EST   30-MAR-2001
DEFINITION nab77a02.x1 Soares_NSF_F8_9W_OT_PA_P_S1 Homo sapiens cDNA clone
            IMAGE:3273507 3' similar to TR:Q9Y5S1 Q9Y5S1 VANILLOID
            RECEPTOR-LIKE PROTEIN 1. ;contains Alu repetitive element;, mRNA
            sequence.
ACCESSION   BF436096
VERSION   BF436096.1 GI:11448411
KEYWORDS   EST.

```
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 513)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -40UP from Gibco
            High quality sequence stop: 423.
FEATURES             Location/Qualifiers
     source          1..513
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3273507"
                     /clone_lib="Soares_NSF_F8_9W_OT_PA_P_S1"
                     /note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with
                     a modified polylinker; Site_1: Not I; Site_2: Eco RI;
                     Equal amounts of plasmid DNA from five normalized
                     libraries were mixed, and ss circles were made in vitro.
                     Following HAP purification, this DNA was used as tracer in
                     a subtractive hybridization reaction. The driver was
                     PCR-amplified cDNAs from pools of 5,000 clones made from
                     the same 5 libraries. The pools consisted of the following
                     libraries and cloneIDs: Soares NbHSF pool 1:
                     30384-310919, 323208-325895 Soares Nb2HP pool 1:
                     145032-147335, 147720-148103, 148872-149255, 15002 -
                     150407, 151176-152327 Soares Nb2HF8-9w pool 1:
                     758280-760583, 772104-774407 Soares NbHPA pool 1:
                     304776-306311, 320136-322823, 326280-326663 Soares NbHOT
                     pool 1: 723720-726407, 739080-740999 Subtraction by Bento
                     Soares and M. Fatima Bonaldo."
BASE COUNT      97 a    142 c    146 g    128 t
ORIGIN
```

```
Query Match             18.3%;  Score 513;  DB 148;  Length 513;
Best Local Similarity   100.0%;  Pred. No. 4.8e-106;
Matches   513;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy     361 atgacctcacctccagctctccagtcttcaggttggagacattagatggaggccaagaa     420
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     513 ATGACCTCACCTCCAGCTCTCCAGTCTTTCAGGTTGGAGACATTAGATGGAGGCCAAGAA   454

Qy     421 gatggctctgaggcggacagaggaaagctggatttgggaggcgggctgctcccatggag     480
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     453 GATGGCTCTGAGGCGGACAGAGGAAAGCTGGATTTTTGGGAGGCGGGGCTGCTCCCATGGAG   394

Qy     481 tcacagttccaggggcggaggaccggaaattcgcccctcagataaggtcaacctcaactac     540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     393 TCACAGTTCCAGGGGCGGAGGACCGGAAATTCGCCCCTCAGATAAGGTCAACCTCAACTAC   334

Qy     541 cgaaagggaacaaggtgccagtcggatccaaaaccgatttgacgagatcggctctc     600
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     333 CGAAAGGGAACAAGGTGCCAGTCGGATCCAAAACCGATTTGACGAGATCGGCTCTTC   274

Qy     601 aatgcggtctcccgggggtgtccccgagggatctgtgctggacttccagagtactgagcaag     660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     273 AATGCGGTCTCCCGGGGTGTCCCCGAGGGATCTGTGCTGGACTTCCAGAGTACCTGAGCAAG   214

Qy     661 accagcaagtacctcaccgactcggaatacacaagaggggctccacaggtagacgtgcctg     720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     213 ACCAGCAAGTACCTCACCGACTCGGAATACACAGAGGGGCTCCACAGGTAGACGTGCCTG   154

Qy     721 atgaaggctgtgctgaaccttaaggacggagtcaatcgcctgaccattctgcacactgctg     780
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     153 ATGAAGGCTGTGCTGAACCTTAAGGACGGAGTCAATCGCCTGCCATTCTGCCACTGCTGCAG    94
```

```
Qy     781 atcgacagggactctggcaatctcagccccctggtaaatgccagtgcacagatgactat     840
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      93 ATCGACAGGGACTCTGGCAATCCTCAGCCCCTGGTAAATGCCCAGTGCCACAGATGACTAT   34

Qy     841 taccgaggcacagcgctctgacactgccatt     873
           ||||||||||||||||||||||||||||||| 1
Db      33 TACCGAGGCCCACACAGCGCTCTGCACATCGCCATT    1
```

```
Search completed: July 21, 2001, 11:11:46
Job time: 4197 sec
```